Journal of Medical Virology 68:417–423 (2002)

# Comparisons of HPV DNA Detection by MY09/11 PCR Methods

**Philip E. Castle,[1]* Mark Schiffman,[1] Patti E. Gravitt,[1,2] Hortense Kendall,[3] Stacy Fishman,[3] Huali Dong,[3] Allan Hildesheim,[1] Rolando Herrero,[4] M. Concepcion Bratti,[4] Mark E. Sherman,[1] Attila Lorincz,[5] John E. Schussler,[6] and Robert D. Burk[3]**

[1]*Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland*
[2]*Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland*
[3]*Albert Einstein College of Medicine, Bronx, New York*
[4]*Proyecto Epidemiologico, Guanacaste, Costa Rica*
[5]*Digene Corporation, Gaithersburg, Maryland*
[6]*Information Management Services, Silver Spring, Maryland*

Two modifications to the original L1 consensus primer human papillomavirus (HPV) PCR method, MY09–MY011, using AmpliTaq DNA polymerase (MY-Taq), were evaluated for HPV DNA detection on clinical specimens from a cohort study of cervical cancer in Costa Rica. First, HPV DNA testing of 2978 clinical specimens by MY09–MY011 primer set, using AmpliTaq Gold DNA polymerase (MY-Gold) were compared with MY-Taq testing. There was 86.8% total agreement ($\kappa = 0.72$, 95%CI = 0.70–75) and 69.6% agreement among positives between MY-Gold and MY-Taq. MY-Gold detected 38% more HPV infections ($P < 0.0001$) and 45% more cancer-associated (high-risk) HPV types ($P < 0.0001$) than MY-Taq, including 12 of the 13 high-risk HPV types. Analyses of discordant results using cytologic diagnoses and detection of HPV DNA by the Hybrid Capture 2 Test suggested that MY-Gold preferentially detected DNA positive specimens with lower HPV viral loads compared with MY-Taq. In a separate analysis, PGMY09–PGMY11 (PGMY-Gold), a redesigned MY09/11 primer set, was compared with MY-Gold for HPV DNA detection (n = 439). There was very good agreement between the two methods ($\kappa = 0.83$; 95%CI = 0.77–0.88) and surprisingly no significant differences in HPV detection ($P = 0.41$). In conclusion, we found MY-Gold to be a more sensitive assay for the detection of HPV DNA than MY-Taq. Our data also suggest that studies reporting HPV DNA detection by PCR need to report the type of polymerase used, as well as other assay specifics, and underscore the need for worldwide standards of testing. ***J. Med. Virol. 68:417–423, 2002.*** © 2002 Wiley-Liss, Inc.

**KEY WORDS:** HPV; MY09/11 PCR; PGMY09/11 PCR

## INTRODUCTION

Although a causal link between human papillomavirus (HPV) infection and cervical cancer has been established [Bosch et al., 1995; Walboomers et al., 1999], improving the detection of HPV DNA by polymerase chain reaction (PCR) methods remains critical to understanding the natural history of HPV and cervical cancer. First, increased accuracy for HPV detection will result in better estimates of absolute risk for cervical cancer and pre-cancerous lesions. Second, methods with even modestly limited sensitivity of HPV detection may misclassify important viral infection patterns, including measurements of viral acquisition/persistence/clearance [Hildesheim et al., 1994; Franco et al., 1999] caused by inherent biologic fluctuations in viral load and differences in specimen quality. Understanding patterns of viral infection is important in clarifying the natural history of HPV infection and perhaps in the use of HPV testing in clinical management of cervical neoplasia. Finally, although HPV infection is the primary cause of cervical cancer, it has become apparent that the vast majority of infections resolve with only mild or no cytologic abnormalities. Thus, natural history studies have been recently directed towards identifying factors that influence the likelihood of an HPV infection progressing to pre-cancer and cancer or being eliminated or suppressed. Misclassification of HPV status or HPV genotype can influence these kinds of analyses and

Grant sponsor: Public Health Service; Grant numbers: NOICP21081, NOICP31061.

*Correspondence to: Dr. Philip E. Castle, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd, MSC 7234, Bethesda, MD 20892.
E-mail: pc95p@nih.gov

Accepted 29 March 2002
DOI 10.1002/jmv.10220
Published online in Wiley InterScience (www.interscience.wiley.com)

results in artifacts, emphasizing the importance of improved HPV detection [Franco, 1992; Schiffman and Schatzkin, 1994].

A number of changes in HPV DNA detection in MY09/MY11 L1 consensus primer PCR have occurred since its introduction during the late 1980s [Bauer et al., 1992]. A primer, HMB01, was first added for improved detection of cancer-associated HPV 51 [Hildesheim et al., 1994]. Using this improved MY09/11 assay (referred to as MY09/11 in the present study), we found approximately 95% of all prevalent cases of cervical intraepithelial neoplasia grade 3 (CIN3) and cancer to be HPV DNA positive in Costa Rica (M. Schiffman, unpublished results). Recently, a new polymerase with better enzymatic characteristics for PCR reactions, AmpliTaq Gold (TaqGold; Perkin-Elmer-Cetus, Norwalk, CT), has been introduced into the MY09/11 assay. In addition, MY09/11 primers were redesigned to minimize base-pair mismatches that result in poor amplification of some HPV types. This new primer system, PGMY09/PGMY11 (PGMY) [Gravitt et al., 2000], has exhibited increased assay performance compared with its predecessor [Gravitt et al., 2000; Kornegay et al., 2001; Coutlee et al., 2002].

To understand the impact of this recent generation of assay changes, we evaluated the analytic performance of these two PCR test modifications, TaqGold and PGMY, to the original MY09/11 system. We first compared HPV detection by MY09/11 using TaqGold (MY-Gold) with the MY09/11 using AmpliTaq DNA polymerase (Perkin-Elmer-Cetus) (MY-Taq). We then compared HPV DNA detection by PGMY, using TaqGold (PGMY-Gold) to MY-Gold.

## MATERIALS AND METHODS

### Population

A natural history study of HPV and cervical cancer was conducted within the National Cancer Institute (NCI)-sponsored, NCI- and local Institutional Review Board-approved population-based cohort of women in Guanacaste, Costa Rica [Herrero et al., 1997, 2000]. Briefly, 10,049 consenting women of the 11,742 women identified in a door-to-door survey residing in Guanacaste agreed to visit one of our study clinics and participated in the enrollment interview. Pelvic examinations were carried out on 9,203 women after excluding virgins ($n = 583$) and those women who were unwilling or unable to undergo an examination ($n = 291$) and including 28 (of 31) supplemental cancer cases identified from major centers to which Guanacaste residents are referred for diagnosis and treatment.

### Data and Specimen Collection

At enrollment, sexually active women who gave consent underwent a pelvic examination, at which time Papanicolau (Pap) smears were prepared, cells were collected for semi-automated ThinPrep cytology (Cytyc, Boxsborough, MA), and Cervigrams (National Testing Laboratories, Fenton, MO) were taken. An additional cervical cell specimen was obtained with a Dacron swab, which was placed in 1.0 ml of specimen transport medium (STM; Digene, Gaithersburg, MD) and stored frozen until use in HPV DNA testing [Herrero et al., 1997, 2000].

All cervical abnormalities identified by visual inspection, cytology, or cervicography were referred to colposcopy. Visible lesions identified by colposcopy were biopsied. A final enrollment diagnosis was assigned to each woman based on a review of cytology, cervigram, and histology: 40 women had invasive cancer, 128 women had high-grade intraepithelial lesions (HSIL), 189 had low-grade intraepithelial lesions lesions (LSIL), 661 had equivocal lesions (ASCUS), and 7,564 were judged normal [Herrero et al., 2000].

### Specimen Selections

***Comparison of MY-Gold and MY-Taq.*** We tested 9,181 of 9,203 enrollment cervical specimens (22 specimens had insufficient material for testing) for HPV DNA using MY-Gold, which resulted in 9,148 valid tests (PCR tests with a positive amplification result, using human β-globin primers indicated an adequate specimen for PCR testing and the test was considered valid). We had previously tested a convenience sample of 3,063 specimens by MY-Taq, resulting in 3,013 valid tests, including 2,974 valid tests as previous reported [Herrero et al., 2000]. A set of 2,978 specimens had valid test results for both MY-Gold and MY-Taq and was used for comparing the two assays. (Note that HPV DNA positivity in this study does not represent population prevalence of HPV because it is not a random sample of the population.)

***Comparison of PGMY-Gold and MY-Gold.*** We randomly selected 244 HPV DNA negatives and 238 HPV DNA positives ($n = 482$) as determined by the original MY-Taq test [Herrero et al., 2000], and masked to previous PCR result. We tested them by PGMY, using TaqGold (PGMY-Gold) and compared them with MY-Gold results described in the previous section. Because of inadequate available material for some specimens, only 453 MY-Gold tests and 449 PGMY-Gold tests were performed, producing 450 and 441 valid results, respectively, as well as 439 specimens with both tests. Repeat testing by MY-Taq was also carried out and demonstrated excellent interassay reproducibility ($\kappa = 1.0$; 95%CI = 0.99–1.0 for overall HPV DNA positivity).

### HPV DNA Testing

***Sample processing.*** Two drops of blue juice solution (Digene Diagnostics, Silver Spring, MD) were added to the samples, vortexed vigorously for 10 sec, and incubated at 37°C for 2 hr. A 100-μl aliquot of this solution was mixed with 400 μl of ammonium acetate/ethanol solution (0.825 M ammonium acetate, 83.5% ethanol); the DNA was precipitated by centrifugation. The DNA pellet was resuspended in 150 μl TE (10 mM Tris, pH 7.5, 0.1 mM EDTA) and stored at −20°C. A blank control (i.e., no sample) was included every 20 samples.

***PCR amplifications.*** HPV DNA detection by MY-Taq was reported previously [Qu et al., 1997]. Two batches of MY09/MY11 primers were synthesized and combined to minimize batch-to-batch variations. The composition of the PCR reactions was the same for MY-Gold as those for MY-Taq, except that 7.5 U TaqGold was used instead of 2.5 U AmpliTaq DNA polymerase. Amplification was performed with 5 μl of the DNA solution from the processed STM specimen. Thermocyling conditions for TaqGold reactions differed from those use for AmpliTaq and included initial denaturation at 95°C for 9 min; thereafter, each cycle consisted of 95°C for 60 sec, 55°C annealing for 60 sec, and extension at 72°C for 60 sec for a total 40 cycles with a final extension at 72°C for 5 min. A 100-cell copy and 2-cell copy SiHa DNA-positive control, and a 100-cell copy of HuH7 DNA-negative control were used per 48 specimens tested. The PGMY primer set was synthesized according to previously published specifications [Gravitt et al., 2000].

***Detection of PCR products.*** Detection of PCR products by gel electrophoresis, Southern blot transfer and hybridization with radiolabeled probes for HPV types 11, 16, 18, and 51 (generic probe) has been described previously [Burk et al., 1996].

***Dot blot procedures.*** Dot blot hybridization was conducted as described [Qu et al., 1997]. All PCR products were hybridized with type-specific probes for HPV types: 6, 11, 16, 18, 26, 31–33, 35, 39, 40, 45, 51–56, 58, 59, 61, 66–68, 70, 71 (AE8), 72, 73, 81 (AE7), 83 (PAP291), 84 (PAP155), 85 (AE5), AE2 (IS39), and AE6. PCR products from the first batch (n = 957) specimens tested by MY-Taq were also hybridized with type-specific probes for HPV 34, 42, 62, 64, 69, and 82 (W13B). Subsequent batches of MY-Taq and MY-Gold PCR products were hybridized with type-specific probes for HPV types 2, 13, 34, 42–44, 57, 62, 64, 69, 82 (W13B), AE9, and AE10, either individually or in combination with other probes. Probes for HPV types 2, 13, 34, 42–44, 57, 62, 64, 69, 74, 82 (W13B), and AE9 were combined in dot blot hybridizations for detection of rare nononcogenic types (dbmix); comparisons of dbmix and untyped HPV DNA positives were restricted to 2,021 specimens, excluding the first batch of specimens for which only a portion of the HPV type probe in the dbmix were included. More than two probes were used to detect each HPV type except for HPV 74, AE5, and AE9. Probe sequences have been previously reported [Qu et al., 1997] except for AE5, AE6, AE7, HPV 71, AE9 and AE10. (The oligonucleotide probe sequences for these probes are as follows: AE5, CTGCAACTACTAATCCAGTTCC; AE6, CCACAGAATACAGTTCTACACGCT and CCCCAGAACCTAAAAAGGAT; AE7, AGCTACATCTGCTGCTGCA and AGAGGACCCTTATGCCGACA; HPV71, CTGTGCTACCAAAACTGTTGAG and AAACAGTCCTCCTCCTGCA; AE9, CAGCAACCTCGCAGGATACG; AE10, CCTACTACACAATCCCCTCCTGCTGCTACAT and ACACAATCCCCTCCTGCTG.) Three experienced investigators interpreted each result, and discrepancies were resolved by consensus.

***Hybrid capture testing.*** As an independent measure of HPV status and viral load for evaluating discordant MY-Taq and MY-Gold results, results from an unbiased subset of specimens (n = 954) tested by the Hybrid Capture 2 Test, using probe B (HC2-B; Digene), a Food and Drug Administration (FDA)-approved diagnostic for HPV DNA previously performed on a stratified subsample of the cohort (n = 1,119) [Schiffman et al., 2000], were used. HC2-B uses RNA probes to detect high-risk HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68). HC2-B was carried out at Digene, according to the manufacturer's specifications. The ratio of the continuous outcome compared with the positive control (1 pg HPV DNA), relative light units per positive control (RLU/PC), is a semiquantitative measure of viral load [Schiffman et al., 2000]. (Note: 1 RLU/PC is the optimized cutoff point for a positive HC2-B test.)

## Statistical Analysis

Women were classified as HPV DNA positive or HPV DNA negative. Test results were ranked hierarchically according to one of three risk groups for cervical cancer: (1) positive for cancer-associated (high-risk) HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) (2); (2) negative for high-risk HPV types but positive for low cancer risk (low-risk) HPV types (including all other HPV types and HPV DNA positives by the generic probe but uncharacterized type-specific probes); or (3) HPV DNA negative (high-risk > low-risk > negative). Analysis of discordant results for MY-Taq and MY-Gold, using HC2-B, were restricted to the 13 high-risk types tested for by HC2-B.

In this study, κ statistics were calculated to evaluate the agreement between methods beyond the agreement expected by chance alone. To evaluate agreement among HPV risk groups, both κ and linear-weighted κ statistics were used. Linear-weighted κ statistics linearly weights disagreement inversely proportional to number of categories of difference between the two assays (i.e., exact agreement is given a weight of 1, a difference of one category (e.g., high-risk vs low-risk HPV types) is given a weight of 0.5, and a difference of 2 categories is given a weight of 0), were used [Landis and Koch, 1977]. McNemar's $\chi^2$ was used to test the significance for unequal distribution of discordant results for matched pair data with dichotomous outcomes and the $\chi^2$ test for symmetry was used to test for marginal homogeneity when comparing multiple categorical outcomes.

Pearson's $\chi^2$ was used to evaluate the differences in the distribution of discordant pairs with cytologic diagnoses. A nonparametric analysis of variance, the Kruskal-Wallis test, was used to evaluate differences between the HC2-B RLU/PC value, a proxy for viral load, and paired test results from MY-Taq and MY-Gold comparison; multiple group comparisons were performed according the Dunn's method [Dunn, 1964].

## RESULTS

### MY-Gold Versus MY-Taq

For 2,978 specimens with valid PCR tests by both MY-Taq and MY-Gold, overall agreement of 86.8%, an agreement among positives of 69.6%, and a κ value of 0.72 (95%CI = 0.70–75) for HPV DNA detection were found. There was an increase in HPV DNA detection by MY-Gold (42.6%) compared with MY-Taq (30.9%). Of the 392 specimens discordant for the HPV DNA result, 371 were positive for MY-Gold and 21 were positive for MY-Taq ($P < 0.0001$, McNemar's $\chi^2$). Of those specimens positive for HPV DNA by both methods, MY-Gold was more likely to detect multiple infections ($P < 0.0001$, symmetry $\chi^2$) (data not shown).

Comparing HPV DNA detection stratified by HPV risk groups (Table I), the overall percentage agreement was 81.5% with a κ value of 0.65 (95%CI = 0.6–68) and a linear weighted κ value of 0.72 (95%CI = 0.69–0.74). MY-Gold detected 45% more high-risk HPV infections than MY-Taq. Furthermore, MY-Gold was more likely than MY-Taq to reclassify specimens into the high-risk group from the non-high-risk group ($P < 0.0001$, symmetry $\chi^2$).

For individual HPV types, MY-Gold detected more infections by 24 of the 35 individual HPV types, including 12 of the 13 high-risk HPV types ($P < 0.05$, McNemar's $\chi^2$) (Table II). Of 129 specimens that were HPV DNA positive but untyped in AmpliTaq test, 116 were typed, 6 were HPV DNA positive but untyped, and 7 were HPV DNA negative by TaqGold. For 46 specimens untyped by MY-Gold (minus the 6 untyped by both methods), 6 were typed and 40 were HPV DNA negative by MY-Taq.

Cytologic diagnoses (Table III) and HC2-B results on a subset of specimens (Fig. 1) were used evaluate discordant PCR test results. Table III demonstrates only moderate concordance for HPV DNA positivity among cytologic normals (62.7%) and ASCUS diagnoses (65.6%). There were more MY-Taq$^-$/MY-Gold$^+$ discordants than MY-Taq$^+$/MY-Gold$^-$ discordants among cytologic normals (36.0% vs 1.3%) and among ASCUS cytology (32.3% vs 2.1%) ($P < 0.0001$, Pearson $\chi^2$). By contrast, there was a high concordance for HPV DNA positivity by both methods for HSIL (93.2%) and cancer (90.9%). Of the 2978 specimens that also had HC2-B testing (n = 954; 32%), the median value of the HC2-B for MY-Taq$^+$/MY-Gold$^-$ discordants (14.4 RLU/PC) (n = 17) was greater than the median value for MY-Taq$^-$/MY-Gold$^+$ discordants (0.61 RLU/PC) (n = 78) ($P < 0.001$) (Figure 1).

TABLE I. Comparison HPV Risk Group Assignments by MY-Taq and MY-Gold*

| | MY-Gold | | | |
|---|---|---|---|---|
| MY-Taq | HPV− | HPV+ low-risk | HPV+ high-risk | Total |
| HPV− | 1,688 | 237 | 134 | 2,059 |
| HPV+, low-risk | 14 | 229 | 137 | 380 |
| HPV+, high-risk | 7 | 22 | 510 | 539 |
| Total: | 1,709 | 488 | 781 | 2,978 |

HPV, human papillomavirus.
*κ = 0.65 (0.63–0.68); linear-weighted κ = 0.72 (0.69–0.74); $P < 0.0001$, symmetry $\chi^2$; 81.5% total agreement.

TABLE II. Comparison of HPV-Type Specific Results for MY-Taq and MY-Gold*

| | MY-Taq n (%) | MY-Gold n (%) | % Positive agreement[a] | $P$[b] |
|---|---|---|---|---|
| 16 | 155 (5.2) | 205 (6.9) | 67.4 | <0.0001 |
| 18 | 46 (1.5) | 81 (2.7) | 46.0 | <0.0001 |
| 31 | 58 (2.0) | 88 (3.0) | 57.0 | <0.0001 |
| 33 | 28 (0.9) | 44 (1.5) | 50.0 | 0.001 |
| 35 | 18 (0.6) | 37 (1.2) | 25.0 | 0.0009 |
| 39 | 30 (1.0) | 67 (2.3) | 34.7 | <0.0001 |
| 45 | 31 (1.0) | 61 (2.1) | 39.4 | <0.0001 |
| 51 | 72 (2.4) | 107 (3.6) | 58.4 | <0.0001 |
| 52 | 70 (2.4) | 97 (3.3) | 56.1 | <0.0001 |
| 56 | 42 (1.4) | 69 (2.3) | 48.0 | <0.0001 |
| 58 | 87 (2.9) | 118 (4.0) | 62.7 | <0.0001 |
| 59 | 12 (0.4) | 20 (0.7) | 33.3 | 0.05 |
| 68 | 19 (0.6) | 23 (0.8) | 55.6 | 0.2 |
| 6 | 32 (1.1) | 34 (1.1) | 40.4 | 0.7 |
| 11 | 13 (0.4) | 18 (0.6) | 29.2 | 0.2 |
| 26 | 9 (0.3) | 9 (0.3) | 0.0 | 1.0 |
| 32 | 9 (0.3) | 18 (0.6) | 12.5 | 0.05 |
| 40 | 8 (0.3) | 10 (0.3) | 38.4 | 0.5 |
| 53 | 71 (2.4) | 116 (3.9) | 58.5 | <0.0001 |
| 54 | 14 (0.5) | 28 (0.9) | 35.5 | 0.002 |
| 55 | 10 (0.3) | 19 (0.6) | 11.5 | 0.06 |
| 61 | 39 (1.3) | 99 (3.3) | 27.8 | <0.0001 |
| 66 | 19 (0.7) | 50 (1.7) | 38.0 | <0.0001 |
| 67 | 13 (0.4) | 9 (0.3) | 15.8 | 0.3 |
| 70 | 53 (1.8) | 111 (3.7) | 46.4 | <0.0001 |
| 71 | 51 (1.7) | 110 (3.7) | 37.6 | <0.0001 |
| 72 | 12 (0.4) | 17 (0.6) | 21.7 | 0.2 |
| 73 | 20 (0.7) | 30 (1.0) | 51.5 | 0.01 |
| 81 | 31 (1.0) | 61 (2.0) | 33.3 | <0.0001 |
| 83 | 20 (0.7) | 48 (1.6) | 33.3 | <0.0001 |
| 84 | 18 (0.6) | 42 (1.4) | 20.0 | 0.0001 |
| 85 | 18 (0.6) | 26 (0.9) | 29.4 | 0.1 |
| AE2 | 12 (0.4) | 25 (0.8) | 37.0 | 0.002 |
| AE6 | 18 (0.6) | 14 (0.5) | 28.0 | 0.4 |
| AE10 | 6 (0.3) | 5 (0.3) | 0.0 | 0.8 |
| dbmix[c,d] | 13 (0.6) | 72 (3.6) | 9.0 | <0.0001 |
| Untyped[c,e] | 129 (6.4) | 52 (2.6) | 3.3 | <0.0001 |

HPV, human papillomavirus.
*n = 2,978.
[a]% Positive agreement (agreement among positives) is the percentage of the total number of positives by either method that are positive by both methods.
[b]McNemar's $\chi^2$ values indicating which method demonstrated statistically significant greater HPV detection ($P < 0.05$).
[c]n = 2,021.
[d]HPV types 2, 13, 34, 42–44, 57, 62, 64, 69, 74, W13B (82), and AE9.
[e]HPV positive by the general probe, but untyped.

### MY-Gold vs PGMY-Gold

In our study of MY-Gold and PGMY-Gold, a high degree of concordance was found (κ = 0.83, 95%CI = 0.77–0.88; 91.6% total agreement; 86.6% agreement among positives) and no significant difference in performance for overall HPV DNA detection ($P = 0.41$, McNemar's chi-square). Likewise, there was very good agreement between the two primer systems (κ = 0.83,

TABLE III. Comparison of Cytologic Diagnoses Versus HPV DNA Positivity by Only MY-Taq, by Only MY-Gold, or by Both Methods

| MY-Taq: / MY-Gold: | Pos / Neg | Neg / Pos | Pos / Pos | Total |
|---|---|---|---|---|
| Normal | 9 (1.3%) | 253 (36.0%) | 440 (62.7%) | 702 |
| ASCUS | 6 (2.1%) | 92 (32.3%) | 187 (65.6%) | 285 |
| LSIL | 2 (1.3%) | 19 (13.1%) | 132 (86.3%) | 153 |
| HSIL | 4 (3.4%) | 4 (3.4%) | 109 (93.2%) | 117 |
| Cancer | 0 (0.0%) | 3 (9.1%) | 30 (90.9%) | 33 |
| Total: | 21 | 371 | 898 | 1,290[a] |

HPV, human papillomavirus; LSIL, low-grade intraepithelial lesion; HSIL, high-grade intraepithelial lesion.
[a]Read numbers across and down, for total.

95%CI 0.78–0.87; linear-weighted $\kappa = 0.86$, 95%CI = 0.83–0.90) and no significant difference in performance ($P = 0.58$, symmetry chi-square) by risk group specific strata (Table IV). There were some type-specific differences in detection by MY-Gold versus PGMY-Gold. MY-Gold detected more infections by HPV types 6, 16, 51, 53, 58, 61, and PAP 291 infections, whereas PGMY detected more infections by HPV types 40, 52, 56, 59, AE6, and AE10 ($P < 0.05$, Pearson $\chi^2$). There were too few discordant pairs (n = 37), MY-Gold$^+$/PGMY-Gold$^-$ and MY-Gold$^-$/PGMY-Gold$^+$, to meaningfully evaluate distribution of cytologic interpretations among them.

## DISCUSSION

Our results suggest that switching from MY-Taq to MY-Gold significantly increased HPV DNA detection by PCR primer set MY09/11. More importantly, compared with MY-Taq, MY-Gold identified 45% more infections by high-risk HPV types and detected more infections for 12 of the 13 individual high-risk HPV types, including HPV 16 ($P < 0.0001$), the type associated with ~50% of all cervical cancer cases [Bosch et al., 1995]. The excellent inter-assay reproducibility of repeat MY-Taq testing (see Materials and Methods) suggests that the improvement of MY-Gold over MY-Taq cannot be explained by the variability in primer quality [Kornegay et al., 2001]. Variability was minimized as a result of combining two syntheses of MY09/11 primers throughout these studies.

Although many specimens were MY-Taq$^-$/MY-Gold$^+$, few specimens were MY-Taq$^+$/MY-Gold$^-$, and the latter could represent random variability in testing by each method. As shown in Table III, relatively more HPV infections among cytologic normals and equivocal lesions (ASCUS) were detected as MY-Taq$^-$/MY-Gold$^+$ than MY-Taq$^+$/MY-Gold$^-$, suggesting that MY-Gold was more sensitive in detecting the lower viral loads that are commonly associated with normal cytology/histology and transient infections [Schiffman et al., 2000]. This interpretation was confirmed by the analysis

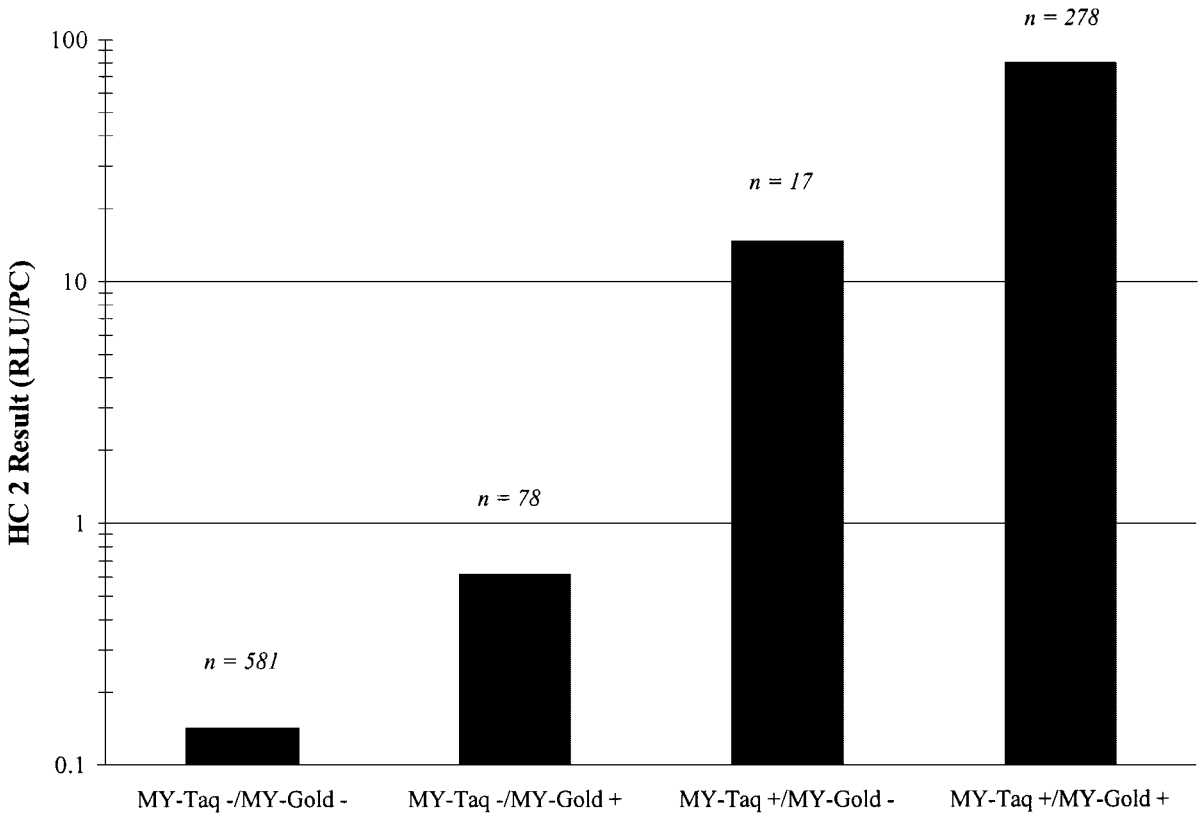


Fig. 1. Comparison of Hybrid Capture 2 test (HC2) results (RLU/PC) versus MY-Taq and MY-Gold results (n = 954) ($P < 0.0001$).

TABLE IV. Agreement in HPV Risk Group Assignment With MY-Gold and PGMY-Gold*

| | PGMY-Gold | | | |
|---|---|---|---|---|
| MY-Gold | HPV− | HPV+, low-risk | HPV+, high-risk | Total |
| HPV− | 162 | 15 | 6 | 183 |
| HPV+, low-risk | 13 | 70 | 7 | 90 |
| HPV+, high-risk | 3 | 4 | 159 | 166 |
| Total: | 178 | 89 | 172 | 439 |

HPV, human papillomavirus.
*$\kappa = 0.83$ (0.78–0.87); linear-weighted $\kappa = 0.86$ (0.83–0.90); $P = 0.58$, symmetry $\chi^2$; 89.1% total agreement.

of HC2-B results for the discordant pairs, which shows that the MY-Taq$^+$/ MY-Gold$^-$ group had a much higher viral load than the MY-Taq$^-$/ MY-Gold$^+$ group. Resolution of genotype by MY-Gold in previous "untyped" MY-Taq$^+$ specimens is also indicative of the sensitivity differences at low viral loads, since the generic probe with more labels is slightly more sensitive than the type-specific oligonucleotide probe.

We offer several explanations for the improved performance of MY-Gold over MY-Taq. First, TaqGold is chemically modified version of AmpliTaq in which a chemical inhibitor is covalently attached to the polymerase. When the PCR reaction reaches the optimal annealing temperatures, the inhibitor is chemically cleaved from the polymerase resulting in active polymerase. Use of TaqGold can prevent nonspecific product formation that can out-compete target-specific amplification and thereby lower analytic assay sensitivity. More importantly, such nonspecific product formation is time dependent (P. Gravitt, unpublished observations), with a greater impact on large population studies in which batch processing is used, as was done in this study. Thus, the seemingly lower amplification efficiency of MY-Taq observed in this study compared with other studies of smaller size [Cope et al., 1997] may be a consequence of scaled-up testing and unavoidable setup time effects. Second, threefold more enzyme was used in the TaqGold PCR reactions than was used in the AmpliTaq PCR reactions and may also account for differences in assay performance. Finally, increased positivity by MY-Gold may be the result of false-positive detection of HPV DNA. Although MY-Gold detected more HPV infections than MY-Taq in this study, it is of note that there were no differences in the detection of HSIL and cancer. In addition, the median RLU/PC value of the HC2-B for the MY-Taq$^-$/MY-Gold$^+$ was below the optimal cutoff point for HPV DNA detection by HC2-B. However, among cytologically normal women, the proportion of women diagnosed with cervical intraepithelial neoplasia grade 2 or more severe at follow-up evaluation did not differ between specimens that were MY-Taq$^-$/MY-Gold$^+$ (26 of 166, 15.7%) and those that were MY-Taq$^+$/MY-Gold$^+$ (36 of 205, 17.6%) ($P = 0.63$). Moreover, the proportion of women who had three or more lifetime sexual partners did not differ significantly between specimens that were MY-Taq$^-$/MY-Gold$^+$ (107 of 268, 39.9%) and those that were MY-Taq$^+$/MY-Gold$^+$ (173 of 509, 34.0%) ($P = 0.10$). Together, these data suggest that the increased detection of HPV DNA by MY-Gold was primarily due to detection of true infections but we cannot discount the possibility of some false-positive detection by MY-Gold.

In contrast with previous reports [Gravitt et al., 2000; Kornegay et al., 2001; Coutlee et al., 2002], we found that PGMY-Gold had a similar analytic sensitivity to MY-Gold for the detection of HPV DNA. Indeed, for several important HPV types, including HPV 16, MY09/11, detected more HPV 16-positive specimens than PGMY whereas the previous studies report greater detection for HPV 16 by PGMY. Similarity in the distribution of discordant pairs among cytologic diagnoses further suggests similar assay sensitivities for two primer systems. These results were surprising because PGMY primers were rationally designed for improved detection of individual HPV types [Gravitt et al., 2000]. These findings warrant caution and need to be reproduced before any meaningful conclusions can be reached.

Finally, suboptimal performance of HPV DNA detection assays (e.g., MY-Taq in the present study) and differences in the relative performance of different primer systems (e.g., PGMY vs MY-Gold in this study compared with previous reports) emphasize the need for worldwide standards of testing that include QC specimens [Nindl et al., 1999]. Indeed, a HPV DNA testing international QC proficiency panel is being developed to address inter-laboratory comparability (Dr. C. Wheeler, personal communication). Without such QA measures, it will be difficult to understand and adjudicate differences in HPV prevalence between natural studies of cervical cancer and also accurately assess the absolute risk associated with HPV infection and attributable risk of HPV cofactors among the HPV infected.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge Julie Buckland (IMS, Rockville, MD) for data management and analytic support. We thank Deidra Kelly (Johns Hopkins University) and Dr. Laurie Mango (Neuromedical Systems, NY) for their collaboration in the interpretation of cytological specimens. Reagents and services were supplied or discounted by Cytyc (Boxsborough, MA, National Testing Laboratories (Fenton, MO), Utah Medical (Midvale, UT), and Neuromedical Systems (Suffern, NY). We offer special recognition for the excellent work of the study staff in Costa Rica, in particular Fernando Cardenas, Manuel Barrantes, Elmer Perez, Lidia Ana Morera, Iris Ugarte, Dr. Mario Alfaro (cytology reading), and Dr. Jorge Morales (colposcopy). We also acknowledge the collaboration of health authorities in Costa Rica for their enthusiastic support of this project and of the outreach workers of the Ministry of Health of Costa Rica who carried out the population census for their dedication to the health of the people of Guanacaste. Dr. Lorincz is the Chief Scientific Officer of Digene and holds Digene stock and stock options. This

work supported by Public Health Services (PHS) Contracts NOICP21081 and NOICP31061 between the National Cancer Institute (NCI), National Institutes of Health, Department of Health and Human Services, and FUCODOCSA (Costa Rican Foundation for Training in Health Sciences), Caja Costarricense de Seguro Social, Costa Rica. R.D. Burk was supported by PHS grant ROICA78527 from the NCI.

## REFERENCES

Bauer HM, Greer CE, Manos MM. 1992. Determination of genital human papillomavirus infection by consensus PCR amplification. In: Herringon CS, McGee JOD, editors. Diagnostic molecular pathology: a practical approach. Oxford: Oxford University Press. p 131–152.

Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV. 1995. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International Biological Study on Cervical Cancer (IBSCC) Study Group. J Natl Cancer Inst 87:796–802.

Burk RD, Ho GY, Beardsley L, Lempa M, Peters M, Bierman R. 1996. Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women. J Infect Dis 174:679–689.

Cope JU, Hildesheim A, Schiffman MH, Manos MM, Lorincz AT, Burk RD, Glass AG, Greer C, Buckland J, Helgesen K, Scott DR, Sherman ME, Kurman RJ, Liaw KL. 1997. Comparison of the hybrid capture tube test and PCR for detection of human papillomavirus DNA in cervical specimens. J Clin Microbiol 35:2262–2265.

Coutlee F, Gravitt P, Kornegay J, Hankins C, Richardson H, Lapointe N, Voyer H, Franco E. 2002. Use of PGMY primers in L1 consensus PCR improves detection of human papillomavirus DNA in genital samples. J Clin Microbiol 40:902–907.

Dunn OJ. 1964. Multiple contrasts using rank sums. Technometrics 6:241–252.

Franco EL. 1992. Measurement errors in epidemiological studies of human papillomavirus and cervical cancer. IARC Sci Publ 119: 181–197.

Franco EL, Villa LL, Sobrinho JP, Prado JM, Rousseau MC, Desy M, Rohan TE. 1999. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. J Infect Dis 180:1415–1423.

Gravitt PE, Peyton CL, Alessi TQ, Wheeler CM, Coutlee F, Hildesheim A, Schiffman MH, Scott DR, Apple RJ. 2000. Improved amplification of genital human papillomaviruses. J Clin Microbiol 38:357–361.

Herrero R, Schiffman MH, Bratti C, Hildesheim A, Balmaceda I, Sherman ME, Greenberg M, Cardenas F, Gomez V, Helgesen K, Morales J, Hutchinson M, Mango L, Alfaro M, Potischman NW, Wacholder S, Swanson C, Brinton LA. 1997. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: The Guanacaste project. Rev Panam Salud Publ 1:362–375.

Herrero R, Hildesheim A, Bratti C, Sherman ME, Hutchinson M, Morales J, Balmaceda I, Greenberg MD, Alfaro M, Burk RD, Wacholder S, Plummer M, Schiffman M. 2000. Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica. J Natl Cancer Inst 92:464–474.

Hildesheim A, Schiffman MH, Gravitt PE, Glass AG, Greer CE, Zhang T, Scott DR, Rush BB, Lawler P, Sherman ME, Kurman RJ, Manos MM. 1994. Persistence of type-specific human papillomavirus infection among cytologically normal women. J Infect Dis 169: 235–240.

Kornegay JR, Shepard AP, Hankins C, Franco E, Lapointe N, Richardson H, Coutlee F. 2001. Nonisotopic detection of human papillomavirus DNA in clinical specimens using a consensus PCR and a generic probe mix in an enzyme-linked immunosorbent assay format. J Clin Microbiol 39:3530–3536.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159–174.

Nindl I, Jacobs M, Walboomers JM, Meijer CJ, Pfister H, Wieland U, Meyer T, Stockfleth E, Klaes R, von Knebel Doeberitz M, Schneider A, Duerst M. 1999. Interlaboratory agreement of different human papillomavirus DNA detection and typing assays in cervical scrapes. Int J Cancer 81:666–668.

Qu W, Jiang G, Cruz Y, Chang CJ, Ho GY, Klein RS, Burk RD. 1997. PCR detection of human papillomavirus: comparison between MY09/MY11 and $GP5^{+}/GP6^{+}$ primer systems. J Clin Microbiol 35: 1304–1310.

Schiffman MH, Schatzkin A. 1994. Test reliability is critically important to molecular epidemiology: an example from studies of human papillomavirus infection and cervical neoplasia. Cancer Res 54:1944s–1947s.

Schiffman M, Herrero R, Hildesheim A, Sherman ME, Bratti M, Wacholder S, Alfaro M, Hutchinson M, Morales J, Greenberg MD, Lorincz AT. 2000. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. JAMA 283:87–93.

Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N. 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 189:12–19.